CHAPTER 3

Sobolev spaces

We will give only the most basic results here. For more information, see Shkoller [**27**], Evans [**8**] (Chapter 5), and Leoni [**20**]. A standard reference is [**1**].

3.1. Weak derivatives

Suppose, as usual, that Ω is an open set in $\mathbb{R}^n$.

Definition 3.1. A function $f \in L^1_{\mathrm{loc}}(\Omega)$ is weakly differentiable with respect to x_i if there exists a function $g_i \in L^1_{\mathrm{loc}}(\Omega)$ such that

$$\int_\Omega f \partial_i \phi \, dx = -\int_\Omega g_i \phi \, dx \qquad \text{for all } \phi \in C_c^\infty(\Omega).$$

The function g_i is called the weak ith partial derivative of f, and is denoted by $\partial_i f$.

Thus, for weak derivatives, the integration by parts formula

$$\int_\Omega f \partial_i \phi \, dx = -\int_\Omega \partial_i f \phi \, dx$$

holds by definition for all $\phi \in C_c^\infty(\Omega)$. Since $C_c^\infty(\Omega)$ is dense in $L^1_{\mathrm{loc}}(\Omega)$, the weak derivative of a function, if it exists, is unique up to pointwise almost everywhere equivalence. Moreover, the weak derivative of a continuously differentiable function agrees with the pointwise derivative. The existence of a weak derivative is, however, *not* equivalent to the existence of a pointwise derivative almost everywhere; see Examples 3.4 and 3.5.

Unless stated otherwise, we will always interpret derivatives as weak derivatives, and we use the same notation for weak derivatives and continuous pointwise derivatives. Higher-order weak derivatives are defined in a similar way.

Definition 3.2. Suppose that $\alpha \in \mathbb{N}_0^n$ is a multi-index. A function $f \in L^1_{\mathrm{loc}}(\Omega)$ has weak derivative $\partial^\alpha f \in L^1_{\mathrm{loc}}(\Omega)$ if

$$\int_\Omega (\partial^\alpha f) \, \phi \, dx = (-1)^{|\alpha|} \int_\Omega f \, (\partial^\alpha \phi) \; dx \qquad \text{for all } \phi \in C_c^\infty(\Omega).$$

3.2. Examples

Let us consider some examples of weak derivatives that illustrate the definition. We denote the weak derivative of a function of a single variable by a prime.

Example 3.3. Define $f \in C(\mathbb{R})$ by

$$f(x) = \begin{cases} x & \text{if } x > 0, \\ 0 & \text{if } x \le 0. \end{cases}$$

We also write $f(x) = x_+$. Then f is weakly differentiable, with

$$f' = \chi_{[0,\infty)}, \tag{3.1}$$

where $\chi_{[0,\infty)}$ is the step function

$$\chi_{[0,\infty)}(x) = \begin{cases} 1 & \text{if } x \geq 0, \\ 0 & \text{if } x < 0. \end{cases}$$

The choice of the value of $f'(x)$ at $x = 0$ is irrelevant, since the weak derivative is only defined up to pointwise almost everwhere equivalence. To prove (3.1), note that for any $\phi \in C_c^\infty(\mathbb{R})$, an integration by parts gives

$$\int f\phi' \, dx = \int_0^\infty x\phi' \, dx = -\int_0^\infty \phi \, dx = -\int \chi_{[0,\infty)}\phi \, dx.$$

Example 3.4. The discontinuous function $f : \mathbb{R} \to \mathbb{R}$

$$f(x) = \begin{cases} 1 & \text{if } x > 0, \\ 0 & \text{if } x < 0. \end{cases}$$

is not weakly differentiable. To prove this, note that for any $\phi \in C_c^\infty(\mathbb{R})$,

$$\int f\phi' \, dx = \int_0^\infty \phi' \, dx = -\phi(0).$$

Thus, the weak derivative $g = f'$ would have to satisfy

$$\int g\phi \, dx = \phi(0) \qquad \text{for all } \phi \in C_c^\infty(\mathbb{R}). \tag{3.2}$$

Assume for contradiction that $g \in L^1_{\mathrm{loc}}(\mathbb{R})$ satisfies (3.2). By considering test functions with $\phi(0) = 0$, we see that g is equal to zero pointwise almost everywhere, and then (3.2) does not hold for test functions with $\phi(0) \neq 0$.

The pointwise derivative of the discontinuous function f in the previous example exists and is zero except at 0, where the function is discontinuous, but the function is not weakly differentiable. The next example shows that even a continuous function that is pointwise differentiable almost everywhere need not have a weak derivative.

Example 3.5. Let $f \in C(\mathbb{R})$ be the Cantor function, which may be constructed as a uniform limit of piecewise constant functions defined on the standard 'middle-thirds' Cantor set C. For example, $f(x) = 1/2$ for $1/3 \leq x \leq 2/3$, $f(x) = 1/4$ for $1/9 \leq x \leq 2/9$, $f(x) = 3/4$ for $7/9 \leq x \leq 8/9$, and so on.[1] Then f is not weakly differentiable. To see this, suppose that $f' = g$ where

$$\int g\phi \, dx = -\int f\phi' \, dx$$

---

[1] The Cantor function is given explicitly by: $f(x) = 0$ if $x \leq 0$; $f(x) = 1$ if $x \geq 1$;

$$f(x) = \frac{1}{2}\sum_{n=1}^\infty \frac{c_n}{2^n}$$

if $x = \sum_{n=1}^\infty c_n/3^n$ with $c_n \in \{0, 2\}$ for all $n \in \mathbb{N}$; and

$$f(x) = \frac{1}{2}\sum_{n=1}^N \frac{c_n}{2^n} + \frac{1}{2^{N+1}}$$

if $x = \sum_{n=1}^\infty c_n/3^n$, with $c_n \in \{0, 2\}$ for $1 \leq n < k$ and $c_k = 1$.

for all test functions ϕ. The complement of the Cantor set in $[0, 1]$ is a union of open intervals,

$$[0,1] \setminus C = \left(\frac{1}{3}, \frac{2}{3}\right) \cup \left(\frac{1}{9}, \frac{2}{9}\right) \cup \left(\frac{7}{9}, \frac{8}{9}\right) \cup \ldots,$$

whose measure is equal to one. Taking test functions ϕ whose supports are compactly contained in one of these intervals, call it I, and using the fact that $f = c_I$ is constant on I, we find that

$$\int g\phi \, dx = -\int_I f\phi' \, dx = -c_I \int_I \phi' \, dx = 0.$$

It follows that $g = 0$ pointwise a.e. on $[0, 1] \setminus C$, and hence if f is weakly differentiable, then $f' = 0$. From the following proposition, however, the only functions with zero weak derivative are the ones that are equivalent to a constant function. This is a contradiction, so the Cantor function is not weakly differentiable.

Proposition 3.6. *If $f : (a, b) \to \mathbb{R}$ is weakly differentiable and $f' = 0$, then f is a constant function.*

Proof. The condition that the weak derivative f' is zero means that

$$\int f\phi' \, dx = 0 \qquad \text{for all } \phi \in C_c^\infty(a, b). \tag{3.3}$$

Choose a fixed test function $\eta \in C_c^\infty(a, b)$ whose integral is equal to one. We may represent an arbitrary test function $\phi \in C_c^\infty(a, b)$ as

$$\phi = A\eta + \psi'$$

where $A \in \mathbb{R}$ and $\psi \in C_c^\infty(a, b)$ are given by

$$A = \int_a^b \phi \, dx, \qquad \psi(x) = \int_a^x [\phi(t) - A\eta(t)] \, dt.$$

Then (3.3) implies that

$$\int f\phi \, dx = A \int f\eta \, dx = c \int \phi \, dx, \qquad c = \int f\eta \, dx.$$

It follows that

$$\int (f - c) \, \phi \, dx = 0 \qquad \text{for all } \phi \in C_c^\infty(a, b),$$

which implies that $f = c$ pointwise almost everywhere, so f is equivalent to a constant function. □

As this discussion illustrates, in defining 'strong' solutions of a differential equation that satisfy the equation pointwise *a.e.*, but which are not necessarily continuously differentiable 'classical' solutions, it is important to include the condition that the solutions are weakly differentiable. For example, up to pointwise *a.e.* equivalence, the only weakly differentiable functions $u : \mathbb{R} \to \mathbb{R}$ that satisfy the ODE

$$u' = 0 \qquad \text{pointwise } a.e.$$

are the constant functions. There are, however, many non-constant functions that are differentiable pointwise *a.e.* and satisfy the ODE pointwise *a.e.*, but these solutions are not weakly differentiable; the step function and the Cantor function are examples.

Example 3.7. For $a \in \mathbb{R}$, define $f : \mathbb{R}^n \to \mathbb{R}$ by

$$f(x) = \frac{1}{|x|^a}. \tag{3.4}$$

Then f is weakly differentiable if $a + 1 < n$ with weak derivative

$$\partial_i f(x) = -\frac{a}{|x|^{a+1}} \frac{x_i}{|x|}.$$

That is, f is weakly differentiable provided that the pointwise derivative, which is defined almost everywhere, is locally integrable. To prove this, suppose $\epsilon > 0$, and let $\phi^\epsilon \in C_c^\infty(\mathbb{R}^n)$ be a cut-off function that is equal to one in $B_\epsilon(0)$ and zero outside $B_{2\epsilon}(0)$. Then

$$f^\epsilon(x) = \frac{1 - \phi^\epsilon(x)}{|x|^a}$$

belongs to $\in C^\infty(\mathbb{R}^n)$ and $f^\epsilon = f$ in $|x| \geq 2\epsilon$. Integrating by parts, we get

$$\int (\partial_i f^\epsilon)\, \phi \, dx = -\int f^\epsilon \, (\partial_i \phi) \, dx.$$

We have

$$\partial_i f^\epsilon(x) = -\frac{a}{|x|^{a+1}} \frac{x_i}{|x|} \left[1 - \phi^\epsilon(x)\right] - \frac{1}{|x|^a} \partial_i \phi^\epsilon(x).$$

Since $|\partial_i \phi^\epsilon| \leq C/\epsilon$ and $|\partial_i \phi^\epsilon| = 0$ when $|x| \leq \epsilon$ or $|x| \geq 2\epsilon$, we have

$$|\partial_i \phi^\epsilon(x)| \leq \frac{C}{|x|}.$$

It follows that

$$|\partial_i f^\epsilon(x)| \leq \frac{C'}{|x|^{a+1}}$$

where C' is a constant independent of ϵ. The result then follows from the dominated convergence theorem.

Alternatively, instead of mollifying f, we can use the truncated function

$$f^\epsilon(x) = \frac{\chi_{B_\epsilon(0)}(x)}{|x|^a}.$$

3.3. Distributions

Although we will not make extensive use of the theory of distributions, it is useful to understand the interpretation of a weak derivative as a distributional derivative. In fact, the definition of the weak derivative by Sobolev, and others, was one motivation for the subsequent development of distribution theory by Schwartz.

Let Ω be an open set in $\mathbb{R}^n$.

Definition 3.8. A sequence $\{\phi_n : n \in \mathbb{N}\}$ of functions $\phi_n \in C_c^\infty(\Omega)$ converges to $\phi \in C_c^\infty(\Omega)$ in the sense of test functions if:

(a) there exists $\Omega' \Subset \Omega$ such that $\operatorname{spt} \phi_n \subset \Omega'$ for every $n \in \mathbb{N}$;

(b) $\partial^\alpha \phi_n \to \partial^\alpha \phi$ as $n \to \infty$ uniformly on Ω for every $\alpha \in \mathbb{N}_0^n$.

The topological vector space $\mathcal{D}(\Omega)$ consists of $C_c^\infty(\Omega)$ equipped with the topology that corresponds to convergence in the sense of test functions.

Note that since the supports of the ϕ_n are contained in the same compactly contained subset, the limit has compact support; and since the derivatives of all orders converge uniformly, the limit is smooth.

The space $\mathcal{D}(\Omega)$ is not metrizable, but it can be shown that the sequential convergence of test functions is sufficient to determine its topology.

A linear functional on $\mathcal{D}(\Omega)$ is a linear map $T : \mathcal{D}(\Omega) \to \mathbb{R}$. We denote the value of T acting on a test function ϕ by $\langle T, \phi\rangle$; thus, T is linear if

$$\langle T, \lambda\phi + \mu\psi\rangle = \lambda\langle T, \phi\rangle + \mu\langle T, \psi\rangle \qquad \text{for all } \lambda, \mu \in \mathbb{R} \text{ and } \phi, \psi \in \mathcal{D}(\Omega).$$

A functional T is continuous if $\phi_n \to \phi$ in the sense of test functions implies that $\langle T, \phi_n\rangle \to \langle T, \phi\rangle$ in $\mathbb{R}$

Definition 3.9. A distribution on Ω is a continuous linear functional

$$T : \mathcal{D}(\Omega) \to \mathbb{R}.$$

A sequence $\{T_n : n \in \mathbb{N}\}$ of distributions converges to T, written $T_n \rightharpoonup T$, if $\langle T_n, \phi\rangle \to \langle T, \phi\rangle$ for every $\phi \in \mathcal{D}(\Omega)$. The topological vector space $\mathcal{D}'(\Omega)$ consists of the distributions on Ω equipped with the topology corresponding to this notion of convergence.

Thus, the space of distributions is the topological dual of the space of test functions.

Example 3.10. The delta-function supported at $a \in \Omega$ is the distribution

$$\delta_a : \mathcal{D}(\Omega) \to \mathbb{R}$$

defined by evaluation of a test function at a:

$$\langle \delta_a, \phi\rangle = \phi(a).$$

This functional is continuous since $\phi_n \to \phi$ in the sense of test functions implies, in particular, that $\phi_n(a) \to \phi(a)$

Example 3.11. Any function $f \in L^1_{\mathrm{loc}}(\Omega)$ defines a distribution $T_f \in \mathcal{D}'(\Omega)$ by

$$\langle T_f, \phi\rangle = \int_\Omega f\phi\, dx.$$

The linear functional T_f is continuous since if $\phi_n \to \phi$ in $\mathcal{D}(\Omega)$, then

$$\sup_{\Omega'} |\phi_n - \phi| \to 0$$

on a set $\Omega' \Subset \Omega$ that contains the supports of the ϕ_n, so

$$|\langle T, \phi_n\rangle - \langle T, \phi\rangle| = \left|\int_{\Omega'} f\,(\phi_n - \phi)\; dx\right| \leq \left(\int_{\Omega'} |f|\, dx\right) \sup_{\Omega'} |\phi_n - \phi| \to 0.$$

Any distribution associated with a locally integrable function in this way is called a regular distribution. We typically regard the function f and the distribution T_f as equivalent.

Example 3.12. If μ is a Radon measure on Ω, then

$$\langle I_\mu, \phi\rangle = \int_\Omega \phi\, d\mu$$

defines a distribution $I_\mu \in \mathcal{D}'(\Omega)$. This distribution is regular if and only if μ is locally absolutely continuous with respect to Lebesgue measure λ, in which case the Radon-Nikodym derivative

$$f = \frac{d\mu}{d\lambda} \in L^1_{\mathrm{loc}}(\Omega)$$

is locally integrable, and

$$\langle I_\mu, \phi \rangle = \int_\Omega f\phi \, dx$$

so $I_\mu = T_f$. On the other hand, if μ is singular with respect to Lebesgue measure (for example, if $\mu = \delta_a$ is the unit point measure supported at $a \in \Omega$), then I_μ is not a regular distribution.

One of the main advantages of distributions is that, in contrast to functions, every distribution is differentiable. The space of distributions may be thought of as the smallest extension of the space of continuous functions that is closed under differentiation.

Definition 3.13. For $1 \le i \le n$, the ith partial derivative of a distribution $T \in \mathcal{D}'(\Omega)$ is the distribution $\partial_i T \in \mathcal{D}'(\Omega)$ defined by

$$\langle \partial_i T, \phi \rangle = -\langle T, \partial_i \phi \rangle \qquad \text{for all } \phi \in \mathcal{D}(\Omega).$$

For $\alpha \in \mathbb{N}_0^n$, the derivative $\partial^\alpha T \in \mathcal{D}'(\Omega)$ of order $|\alpha|$ is defined by

$$\langle \partial^\alpha T, \phi \rangle = (-1)^{|\alpha|} \langle T, \partial^\alpha \phi \rangle \qquad \text{for all } \phi \in \mathcal{D}(\Omega).$$

Note that if $T \in \mathcal{D}'(\Omega)$, then it follows from the linearity and continuity of the derivative $\partial^\alpha : \mathcal{D}(\Omega) \to \mathcal{D}(\Omega)$ on the space of test functions that $\partial^\alpha T$ is a continuous linear functional on $\mathcal{D}(\Omega)$. Thus, $\partial^\alpha T \in \mathcal{D}'(\Omega)$ for any $T \in \mathcal{D}'(\Omega)$. It also follows that the distributional derivative $\partial^\alpha : \mathcal{D}'(\Omega) \to \mathcal{D}'(\Omega)$ is linear and continuous on the space of distributions; in particular if $T_n \rightharpoonup T$, then $\partial^\alpha T_n \rightharpoonup \partial^\alpha T$.

Let $f \in L^1_{\mathrm{loc}}(\Omega)$ be a locally integrable function and $T_f \in \mathcal{D}'(\Omega)$ the associated regular distribution defined in Example 3.11. Suppose that the distributional derivative of T_f is a regular distribution

$$\partial_i T_f = T_{g_i} \qquad g_i \in L^1_{\mathrm{loc}}(\Omega).$$

Then it follows from the definitions that

$$\int_\Omega f \partial_i \phi \, dx = -\int_\Omega g_i \phi \, dx \qquad \text{for all } \phi \in C_c^\infty(\Omega).$$

Thus, Definition 3.1 of the weak derivative may be restated as follows: A locally integrable function is weakly differentiable if its distributional derivative is regular, and its weak derivative is the locally integrable function corresponding to the distributional derivative.

The distributional derivative of a function exists even if the function is not weakly differentiable.

Example 3.14. If f is a function of bounded variation, then the distributional derivative of f is a finite Radon measure, which need not be regular. For example, the distributional derivative of the step function is the delta-function, and the distributional derivative of the Cantor function is the corresponding Lebesgue-Stieltjes measure supported on the Cantor set.

Example 3.15. The derivative of the delta-function δ_a supported at a, defined in Example 3.10, is the distribution $\partial_i \delta_a$ defined by

$$\langle \partial_i \delta_a, \phi \rangle = -\partial_i \phi(a).$$

This distribution is neither regular nor a Radon measure.

Differential equations are typically thought of as equations that relate functions. The use of weak derivatives and distribution theory leads to an alternative point of view of linear differential equations as linear functionals acting on test functions. Using this perspective, given suitable estimates, one can obtain simple and general existence results for weak solutions of linear PDEs by the use of the Hahn-Banach, Riesz representation, or other duality theorems for the existence of bounded linear functionals.

While distribution theory provides an effective general framework for the analysis of linear PDEs, it is less useful for nonlinear PDEs because one cannot define a product of distributions that extends the usual product of smooth functions in an unambiguous way. For example, what is $T_f \delta_a$ if f is a locally integrable function that is discontinuous at a? There are difficulties even for regular distributions. For example, $f : x \mapsto |x|^{-n/2}$ is locally integrable on $\mathbb{R}^n$ but f^2 is not, so how should one define the distribution $(T_f)^2$?

3.4. Properties of weak derivatives

We collect here some properties of weak derivatives. The first result is a product rule.

Proposition 3.16. *If $f \in L^1_{\mathrm{loc}}(\Omega)$ has weak partial derivative $\partial_i f \in L^1_{\mathrm{loc}}(\Omega)$ and $\psi \in C^\infty(\Omega)$, then ψf is weakly differentiable with respect to x_i and*

$$\partial_i(\psi f) = (\partial_i \psi) f + \psi(\partial_i f). \tag{3.5}$$

Proof. Let $\phi \in C_c^\infty(\Omega)$ be any test function. Then $\psi\phi \in C_c^\infty(\Omega)$ and the weak differentiability of f implies that

$$\int_\Omega f \partial_i(\psi\phi)\, dx = -\int_\Omega (\partial_i f)\psi\phi\, dx.$$

Expanding $\partial_i(\psi\phi) = \psi(\partial_i\phi) + (\partial_i\psi)\phi$ in this equation and rearranging the result, we get

$$\int_\Omega \psi f(\partial_i \phi)\, dx = -\int_\Omega \left[(\partial_i\psi) f + \psi(\partial_i f)\right]\phi\, dx \qquad \text{for all } \phi \in C_c^\infty(\Omega).$$

Thus, ψf is weakly differentiable and its weak derivative is given by (3.5). $\square$

The commutativity of weak derivatives follows immediately from the commutativity of derivatives applied to smooth functions.

Proposition 3.17. *Suppose that $f \in L^1_{\mathrm{loc}}(\Omega)$ and that the weak derivatives $\partial^\alpha f$, $\partial^\beta f$ exist for multi-indices $\alpha, \beta \in \mathbb{N}_0^n$. Then if any one of the weak derivatives $\partial^{\alpha+\beta} f$, $\partial^\alpha \partial^\beta f$, $\partial^\beta \partial^\alpha f$ exists, all three derivatives exist and are equal.*

Proof. Using the existence of $\partial^\alpha u$, and the fact that $\partial^\beta \phi \in C_c^\infty(\Omega)$ for any $\phi \in C_c^\infty(\Omega)$, we have

$$\int_\Omega \partial^\alpha u \partial^\beta \phi\, dx = (-1)^{|\alpha|} \int_\Omega u \partial^{\alpha+\beta} \phi\, dx.$$

This equation shows that $\partial^{\alpha+\beta}u$ exists if and only if $\partial^\beta\partial^\alpha u$ exists, and in that case the weak derivatives are equal. Using the same argument with α and β exchanged, we get the result. $\square$

Example 3.18. Consider functions of the form

$$u(x,y) = f(x) + g(y).$$

Then $u \in L^1_{\mathrm{loc}}(\mathbb{R}^2)$ if and only if $f, g \in L^1_{\mathrm{loc}}(\mathbb{R})$. The weak derivative $\partial_x u$ exists if and only if the weak derivative f' exists, and then $\partial_x u(x,y) = f'(x)$. To see this, we use Fubini's theorem to get for any $\phi \in C_c^\infty(\mathbb{R}^2)$ that

$$\begin{aligned}&\int u(x,y)\partial_x\phi(x,y)\,dxdy \\ &\qquad = \int f(x)\partial_x\left[\int \phi(x,y)\,dy\right]dx + \int g(y)\left[\int \partial_x\phi(x,y)\,dx\right]dy.\end{aligned}$$

Since ϕ has compact support,

$$\int \partial_x\phi(x,y)\,dx = 0.$$

Also,

$$\int \phi(x,y)\,dy = \xi(x)$$

is a test function $\xi \in C_c^\infty(\mathbb{R})$. Moreover, by taking $\phi(x,y) = \xi(x)\eta(y)$, where $\eta \in C_c^\infty(\mathbb{R})$ is an arbitrary test function with integral equal to one, we can get every $\xi \in C_c^\infty(\mathbb{R})$. Since

$$\int u(x,y)\partial_x\phi(x,y)\,dxdy = \int f(x)\xi'(x)\,dx,$$

it follows that $\partial_x u$ exists if and only if f' exists, and then $\partial_x u = f'$.

In that case, the mixed derivative $\partial_y\partial_x u$ also exists, and is zero, since using Fubini's theorem as before

$$\int f'(x)\partial_y\phi(x,y)\,dxdy = \int f'(x)\left[\int \partial_y\phi(x,y)\,dy\right]dx = 0.$$

Similarly $\partial_y u$ exists if and only if g' exists, and then $\partial_y u = g'$ and $\partial_x\partial_y u = 0$.

The second-order weak derivative $\partial_{xy}u$ exists without any differentiability assumptions on $f, g \in L^1_{\mathrm{loc}}(\mathbb{R})$ and is equal to zero. For any $\phi \in C_c^\infty(\mathbb{R}^2)$, we have

$$\begin{aligned}&\int u(x,y)\partial_{xy}\phi(x,y)\,dxdy \\ &\qquad = \int f(x)\partial_x\left(\int \partial_y\phi(x,y)\,dy\right)dx + \int g(y)\partial_y\left(\int \partial_x\phi(x,y)\,dx\right)dy \\ &\qquad = 0.\end{aligned}$$

Thus, the mixed derivatives $\partial_x\partial_y u$ and $\partial_y\partial_x u$ are equal, and are equal to the second-order derivative $\partial_{xy}u$, whenever both are defined.

Weak derivatives combine well with mollifiers. If Ω is an open set in $\mathbb{R}^n$ and $\epsilon > 0$, we define Ω^ϵ as in (1.7) and let η^ϵ be the standard mollifier (1.6).

Theorem 3.19. *Suppose that $f \in L^1_{\mathrm{loc}}(\Omega)$ has weak derivative $\partial^\alpha f \in L^1_{\mathrm{loc}}(\Omega)$. Then $\eta^\epsilon * f \in C^\infty(\Omega^\epsilon)$ and*

$$\partial^\alpha \left(\eta^\epsilon * f\right) = \eta^\epsilon * \left(\partial^\alpha f\right).$$

Moreover,

$$\partial^\alpha \left(\eta^\epsilon * f\right) \to \partial^\alpha f \qquad \text{in } L^1_{\mathrm{loc}}(\Omega) \text{ as } \epsilon \to 0^+.$$

Proof. From Theorem 1.28, we have $\eta^\epsilon * f \in C^\infty(\Omega^\epsilon)$ and

$$\partial^\alpha \left(\eta^\epsilon * f\right) = \left(\partial^\alpha \eta^\epsilon\right) * f.$$

Using the fact that $y \mapsto \eta^\epsilon(x-y)$ defines a test function in $C_c^\infty(\Omega)$ for any fixed $x \in \Omega^\epsilon$ and the definition of the weak derivative, we have

$$\begin{aligned}
\left(\partial^\alpha \eta^\epsilon\right) * f(x) &= \int \partial_x^\alpha \eta^\epsilon(x-y) f(y)\, dy \\
&= (-1)^{|\alpha|} \int \partial_y^\alpha \eta^\epsilon(x-y) f(y) \\
&= \int \eta^\epsilon(x-y) \partial^\alpha f(y)\, dy \\
&= \eta^\epsilon * \left(\partial^\alpha f\right)(x)
\end{aligned}$$

Thus $\left(\partial^\alpha \eta^\epsilon\right) * f = \eta^\epsilon * \left(\partial^\alpha f\right)$. Since $\partial^\alpha f \in L^1_{\mathrm{loc}}(\Omega)$, Theorem 1.28 implies that

$$\eta^\epsilon * \left(\partial^\alpha f\right) \to \partial^\alpha f$$

in $L^1_{\mathrm{loc}}(\Omega)$, which proves the result. □

The next result gives an alternative way to characterize weak derivatives as limits of derivatives of smooth functions.

Theorem 3.20. *A function $f \in L^1_{\mathrm{loc}}(\Omega)$ is weakly differentiable in Ω if and only if there is a sequence $\{f_n\}$ of functions $f_n \in C^\infty(\Omega)$ such that $f_n \to f$ and $\partial^\alpha f_n \to g$ in $L^1_{\mathrm{loc}}(\Omega)$. In that case the weak derivative of f is given by $g = \partial^\alpha f \in L^1_{\mathrm{loc}}(\Omega)$.*

Proof. If f is weakly differentiable, we may construct an appropriate sequence by mollification as in Theorem 3.19. Conversely, suppose that such a sequence exists. Note that if $f_n \to f$ in $L^1_{\mathrm{loc}}(\Omega)$ and $\phi \in C_c(\Omega)$, then

$$\int_\Omega f_n \phi\, dx \to \int_\Omega f \phi\, dx \qquad \text{as } n \to \infty,$$

since if $K = \operatorname{spt} \phi \Subset \Omega$

$$\left| \int_\Omega f_n \phi\, dx - \int_\Omega f \phi\, dx \right| = \left| \int_K \left(f_n - f\right) \phi\, dx \right| \leq \sup_K |\phi| \int_K \left|f_n - f\right|\, dx \to 0.$$

Thus, for any $\phi \in C_c^\infty(\Omega)$, the L^1_{loc}-convergence of f_n and $\partial^\alpha f_n$ implies that

$$\begin{aligned}
\int_\Omega f \partial^\alpha \phi\, dx &= \lim_{n\to\infty} \int_\Omega f_n \partial^\alpha \phi\, dx \\
&= (-1)^{|\alpha|} \lim_{n\to\infty} \int_\Omega \partial^\alpha f_n \phi\, dx \\
&= (-1)^{|\alpha|} \int_\Omega g \phi\, dx.
\end{aligned}$$

So f is weakly differentiable and $\partial^\alpha f = g$. □

We can use this approximation result to derive properties of the weak derivative as a limit of corresponding properties of smooth functions. The following weak versions of the product and chain rule, which are not stated in maximum generality, may be derived in this way.

Proposition 3.21. *Let Ω be an open set in $\mathbb{R}^n$.*

(1) *Suppose that $a \in C^1(\Omega)$ and $u \in L^1_{\mathrm{loc}}(\Omega)$ is weakly differentiable. Then au is weakly differentiable and*
$$\partial_i(au) = a\,(\partial_i u) + (\partial_i a)\,u.$$
(2) *Suppose that $f : \mathbb{R} \to \mathbb{R}$ is a continuously differentiable function with $f' \in L^\infty(\mathbb{R})$ bounded, and $u \in L^1_{\mathrm{loc}}(\Omega)$ is weakly differentiable. Then $v = f \circ u$ is weakly differentiable and*
$$\partial_i v = f'(u)\partial_i u.$$
(3) *Suppose that $\phi : \Omega \to \widetilde{\Omega}$ is a C^1-diffeomorphism of Ω onto $\widetilde{\Omega} = \phi(\Omega) \subset \mathbb{R}^n$. For $u \in L^1_{\mathrm{loc}}(\Omega)$, define $v \in L^1_{\mathrm{loc}}(\widetilde{\Omega})$ by $v = u \circ \phi^{-1}$. Then v is weakly differentiable in $\widetilde{\Omega}$ if and only if u is weakly differentiable in Ω, and*
$$\frac{\partial u}{\partial x_i} = \sum_{j=1}^{n} \frac{\partial \phi_j}{\partial x_i}\frac{\partial v}{\partial y_j} \circ \phi.$$

Proof. We prove (2) as an example. Since $f' \in L^\infty$, f is globally Lipschitz and there exists a constant M such that
$$|f(s) - f(t)| \le M|s - t| \qquad \text{for all } s, t \in \mathbb{R}.$$
Choose $u_n \in C^\infty(\Omega)$ such that $u_n \to u$ and $\partial_i u_n \to \partial_i u$ in $L^1_{\mathrm{loc}}(\Omega)$, where $u_n \to u$ pointwise almost everywhere in Ω. Let $v = f \circ u$ and $v_n = f \circ u_n \in C^1(\Omega)$, with
$$\partial_i v_n = f'(u_n)\partial_i u_n \in C(\Omega).$$
If $\Omega' \Subset \Omega$, then
$$\int_{\Omega'} |v_n - v|\,dx = \int_{\Omega'} |f(u_n) - f(u)|\,dx \le M\int_{\Omega'} |u_n - u|\,dx \to 0$$
as $n \to \infty$. Also, we have
$$\begin{aligned}\int_{\Omega'} |\partial_i v_n - f'(u)\partial_i u|\,dx &= \int_{\Omega'} |f'(u_n)\partial_i u_n - f'(u)\partial_i u|\,dx \\ &\le \int_{\Omega'} |f'(u_n)|\,|\partial_i u_n - \partial_i u|\,dx \\ &\quad + \int_{\Omega'} |f'(u_n) - f'(u)|\,|\partial_i u|\,dx.\end{aligned}$$
Then
$$\int_{\Omega'} |f'(u_n)|\,|\partial_i u_n - \partial_i u|\,dx \le M\int_{\Omega'} |\partial_i u_n - \partial_i u|\,dx \to 0.$$
Moreover, since $f'(u_n) \to f'(u)$ pointwise *a.e.*, and
$$|f'(u_n)|\,|\partial_i u_n - \partial_i u| \le 2M\,|\partial_i u|$$
the dominated convergence theorem implies that
$$\int_{\Omega'} |f'(u_n)|\,|\partial_i u_n - \partial_i u|\,dx \to 0 \qquad \text{as } n \to \infty.$$

It follows that $v_n \to f \circ u$ and $\partial_i v_n \to f'(u)\partial_i u$ in L^1_{loc}. Then Theorem 3.20, in which it is sufficient but not necessary that the approximating functions are C^∞, implies that $f \circ u$ is weakly differentiable with the weak derivative stated. $\square$

In fact, (2) remains valid if $f \in W^{1,\infty}(\mathbb{R})$ is globally Lipschitz but not necessarily C^1. We will prove this is the useful special case that $f(u) = |u|$.

Proposition 3.22. *If $u \in L^1_{\mathrm{loc}}(\Omega)$ has the weak derivative $\partial_i u \in L^1_{\mathrm{loc}}(\Omega)$, then $|u| \in L^1_{\mathrm{loc}}(\Omega)$ is weakly differentiable and*

$$\partial_i |u| = \begin{cases} \partial_i u & \text{if } u > 0, \\ 0 & \text{if } u = 0, \\ -\partial_i u & \text{if } u < 0. \end{cases} \tag{3.6}$$

Proof. Let

$$f^\epsilon(t) = \sqrt{t^2 + \epsilon^2}.$$

Since f^ϵ is C^1 and globally Lipschitz, Proposition 3.21 implies that $f^\epsilon(u)$ is weakly differentiable, and for every $\phi \in C_c^\infty(\Omega)$

$$\int_\Omega f^\epsilon(u)\partial_i\phi\, dx = -\int_\Omega \frac{u\partial_i u}{\sqrt{u^2+\epsilon^2}}\phi\, dx.$$

Taking the limit of this equation as $\epsilon \to 0$ and using the dominated convergence theorem, we conclude that

$$\int_\Omega |u|\partial_i\phi\, dx = -\int_\Omega (\partial_i |u|)\phi\, dx$$

where $\partial_i|u|$ is given by (3.6). $\square$

It follows immediately from this result that the positive and negative parts of $u = u^+ - u^-$, given by

$$u^+ = \frac{1}{2}\left(|u| + u\right), \qquad u^- = \frac{1}{2}\left(|u| - u\right),$$

are weakly differentiable if u is weakly differentiable, with

$$\partial_i u^+ = \begin{cases} \partial_i u & \text{if } u > 0, \\ 0 & \text{if } u \le 0, \end{cases} \qquad \partial_i u^- = \begin{cases} 0 & \text{if } u \ge 0, \\ -\partial_i u & \text{if } u < 0, \end{cases}$$

3.5. Sobolev spaces

Sobolev spaces consist of functions whose weak derivatives belong to L^p. These spaces provide one of the most useful settings for the analysis of PDEs.

Definition 3.23. Suppose that Ω is an open set in $\mathbb{R}^n$, $k \in \mathbb{N}$, and $1 \le p \le \infty$. The Sobolev space $W^{k,p}(\Omega)$ consists of all locally integrable functions $f : \Omega \to \mathbb{R}$ such that

$$\partial^\alpha f \in L^p(\Omega) \qquad \text{for } 0 \le |\alpha| \le k.$$

We write $W^{k,2}(\Omega) = H^k(\Omega)$.

The Sobolev space $W^{k,p}(\Omega)$ is a Banach space when equipped with the norm

$$\|f\|_{W^{k,p}(\Omega)} = \left(\sum_{|\alpha| \le k} \int_\Omega |\partial^\alpha f|^p\, dx\right)^{1/p}$$

for $1 \leq p < \infty$ and

$$\|f\|_{W^{k,\infty}(\Omega)} = \max_{|\alpha|\leq k} \sup_{\Omega} |\partial^\alpha f|.$$

As usual, we identify functions that are equal almost everywhere. We will use these norms as the standard ones on $W^{k,p}(\Omega)$, but there are other equivalent norms *e.g.*

$$\|f\|_{W^{k,p}(\Omega)} = \sum_{|\alpha|\leq k} \left(\int_\Omega |\partial^\alpha f|^p \, dx\right)^{1/p},$$

$$\|f\|_{W^{k,p}(\Omega)} = \max_{|\alpha|\leq k} \left(\int_\Omega |\partial^\alpha f|^p \, dx\right)^{1/p}.$$

The space $H^k(\Omega)$ is a Hilbert space with the inner product

$$\langle f, g\rangle = \sum_{|\alpha|\leq k} \int_\Omega (\partial^\alpha f)(\partial^\alpha g) \, dx.$$

We will consider the following properties of Sobolev spaces in the simplest settings.

(1) Approximation of Sobolev functions by smooth functions;
(2) Embedding theorems;
(3) Boundary values of Sobolev functions and trace theorems;
(4) Compactness results.

3.6. Approximation of Sobolev functions

To begin with, we consider Sobolev functions defined on all of $\mathbb{R}^n$. They may be approximated in the Sobolev norm by by test functions.

Theorem 3.24. *For $k \in \mathbb{N}$ and $1 \leq p < \infty$, the space $C_c^\infty(\mathbb{R}^n)$ is dense in $W^{k,p}(\mathbb{R}^n)$*

Proof. Let $\eta^\epsilon \in C_c^\infty(\mathbb{R}^n)$ be the standard mollifier and $f \in W^{k,p}(\mathbb{R}^n)$. Then Theorem 1.28 and Theorem 3.19 imply that $\eta^\epsilon * f \in C^\infty(\mathbb{R}^n) \cap W^{k,p}(\mathbb{R}^n)$ and for $|\alpha| \leq k$

$$\partial^\alpha (\eta^\epsilon * f) = \eta^\epsilon * (\partial^\alpha f) \to \partial^\alpha f \qquad \text{in } L^p(\mathbb{R}^n) \text{ as } \epsilon \to 0^+.$$

It follows that $\eta^\epsilon * f \to f$ in $W^{k,p}(\mathbb{R}^n)$ as $\epsilon \to 0$. Therefore $C^\infty(\mathbb{R}^n) \cap W^{k,p}(\mathbb{R}^n)$ is dense in $W^{k,p}(\mathbb{R}^n)$.

Now suppose that $f \in C^\infty(\mathbb{R}^n) \cap W^{k,p}(\mathbb{R}^n)$, and let $\phi \in C_c^\infty(\mathbb{R}^n)$ be a cut-off function such that

$$\phi(x) = \begin{cases} 1 & \text{if } |x| \leq 1, \\ 0 & \text{if } |x| \geq 2. \end{cases}$$

Define $\phi^R(x) = \phi(x/R)$ and $f^R = \phi^R f \in C_c^\infty(\mathbb{R}^n)$. Then, by the Leibnitz rule,

$$\partial^\alpha f^R = \phi^R \partial^\alpha f + \frac{1}{R} h^R$$

where h^R is bounded in L^p uniformly in R. Hence, by the dominated convergence theorem

$$\partial^\alpha f^R \to \partial^\alpha f \qquad \text{in } L^p \text{ as } R \to \infty,$$

so $f^R \to f$ in $W^{k,p}(\mathbb{R}^n)$ as $R \to \infty$. It follows that $C_c^\infty(\Omega)$ is dense in $W^{k,p}(\mathbb{R}^n)$. $\square$

If Ω is a proper open subset of $\mathbb{R}^n$, then $C_c^\infty(\Omega)$ is not dense in $W^{k,p}(\Omega)$. Instead, its closure is the space of functions $W_0^{k,p}(\Omega)$ that 'vanish on the boundary $\partial\Omega$.' We discuss this further below. The space $C^\infty(\Omega)\cap W^{k,p}(\Omega)$ is dense in $W^{k,p}(\Omega)$ for any open set Ω (Meyers and Serrin, 1964), so that $W^{k,p}(\Omega)$ may alternatively be defined as the completion of the space of smooth functions in Ω whose derivatives of order less than or equal to k belong to $L^p(\Omega)$. Such functions need not extend to continuous functions on $\overline{\Omega}$ or be bounded on Ω.

3.7. Sobolev embedding: $p < n$

> G. H. Hardy reported Harald Bohr as saying 'all analysts spend half their time hunter through the literature for inequalities which they want to use but cannot prove.'[2]

Let us first consider the following basic question: Can we estimate the $L^q(\mathbb{R}^n)$-norm of a smooth, compactly supported function in terms of the $L^p(\mathbb{R}^n)$-norm of its derivative? As we will show, given $1 \le p < n$, this is possible for a unique value of q, called the Sobolev conjugate of p.

We may motivate the answer by means of a scaling argument. We are looking for an estimate of the form

$$\|f\|_{L^q} \le C\|Df\|_{L^p} \qquad \text{for all } f \in C_c^\infty(\mathbb{R}^n) \tag{3.7}$$

for some constant $C = C(p, q, n)$. For $\lambda > 0$, let f_λ denote the rescaled function

$$f_\lambda(x) = f\left(\frac{x}{\lambda}\right).$$

Then, changing variables $x \mapsto \lambda x$ in the integrals that define the L^p, L^q norms, with $1 \le p, q < \infty$, and using the fact that

$$Df_\lambda = \frac{1}{\lambda}(Df)_\lambda$$

we find that

$$\begin{aligned} \left(\int_{\mathbb{R}^n} |Df_\lambda|^p \, dx\right)^{1/p} &= \lambda^{n/p-1}\left(\int_{\mathbb{R}^n} |Df|^p \, dx\right)^{1/p}, \\ \left(\int_{\mathbb{R}^n} |f_\lambda|^q \, dx\right)^{1/q} &= \lambda^{n/q}\left(\int_{\mathbb{R}^n} |f|^q \, dx\right)^{1/q}. \end{aligned}$$

These norms must scale according to the same exponent if we are to have an inequality of the desired form, otherwise we can violate the inequality by taking $\lambda \to 0$ or $\lambda \to \infty$. The equality of exponents implies that $q = p^*$ where p^* satifies

$$\frac{1}{p^*} = \frac{1}{p} - \frac{1}{n}. \tag{3.8}$$

Note that we need $1 \le p < n$ to ensure that $p^* > 0$, in which case $p < p^* < \infty$. We assume that $n \ge 2$. Writing the solution of (3.8) for p^* explicitly, we make the following definition.

Definition 3.25. If $1 \le p < n$, the Sobolev conjugate p^* of p is

$$p^* = \frac{np}{n-p}.$$

[2] From the Introduction of [**13**].

Thus, an estimate of the form (3.7) is possible only if $q = p^*$; we will show that (3.7) is, in fact, true when $q = p^*$. This result was obtained by Sobolev (1938), who used potential-theoretic methods (*c.f.* Section 5.D). The proof we give is due to Nirenberg (1959). The inequality is usually called the Gagliardo-Nirenberg inequality or Sobolev inequality (or Gagliardo-Nirenberg-Sobolev inequality . . .).

Before describing the proof, we introduce some notation, explain the main idea, and establish a preliminary inequality.

For $1 \le i \le n$ and $x = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$, let

$$x_i' = (x_1, \dots, \hat{x}_i, \dots x_n) \in \mathbb{R}^{n-1},$$

where the 'hat' means that the ith coordinate is omitted. We write $x = (x_i, x_i')$ and denote the value of a function $f : \mathbb{R}^n \to \mathbb{R}$ at x by

$$f(x) = f\left(x_i, x_i'\right).$$

We denote the partial derivative with respect to x_i by ∂_i.

If f is smooth with compact support, the fundamental theorem of calculus implies that

$$f(x) = \int_{-\infty}^{x_i} \partial_i f(t, x_i')\, dt.$$

Taking absolute values, we get

$$|f(x)| \le \int_{-\infty}^{\infty} \left|\partial_i f(t, x_i')\right|\, dt.$$

We can improve the constant in this estimate by using the fact that

$$\int_{-\infty}^{\infty} \partial_i f(t, x_i')\, dt = 0.$$

Lemma 3.26. *Suppose that $g : \mathbb{R} \to \mathbb{R}$ is an integrable function with compact support such that $\int g\, dt = 0$. If*

$$f(x) = \int_{-\infty}^{x} g(t)\, dt,$$

then

$$|f(x)| \le \frac{1}{2} \int |g|\, dt.$$

Proof. Let $g = g_+ - g_-$ where the nonnegative functions g_+, g_- are defined by $g_+ = \max(g, 0)$, $g_- = \max(-g, 0)$. Then $|g| = g_+ + g_-$ and

$$\int g_+\, dt = \int g_-\, dt = \frac{1}{2} \int |g|\, dt.$$

It follows that

$$f(x) \le \int_{-\infty}^{x} g_+(t)\, dt \le \int_{-\infty}^{\infty} g_+(t)\, dt \le \frac{1}{2} \int |g|\, dt,$$

$$f(x) \ge -\int_{-\infty}^{x} g_-(t)\, dt \ge -\int_{-\infty}^{\infty} g_-(t)\, dt \ge -\frac{1}{2} \int |g|\, dt,$$

which proves the result. □

Thus, for $1 \leq i \leq n$ we have

$$|f(x)| \leq \frac{1}{2} \int_{-\infty}^{\infty} |\partial_i f(t, x_i')| \, dt.$$

The idea of the proof is to average a suitable power of this inequality over the i-directions and integrate the result to estimate f in terms of Df. In order to do this, we use the following inequality, which estimates the L^1-norm of a function of $x \in \mathbb{R}^n$ in terms of the L^{n-1}-norms of n functions of $x_i' \in \mathbb{R}^{n-1}$ whose product bounds the original function pointwise.

Theorem 3.27. *Suppose that $n \geq 2$ and*

$$\left\{g_i \in C_c^\infty(\mathbb{R}^{n-1}) : 1 \leq i \leq n\right\}$$

are nonnegative functions. Define $g \in C_c^\infty(\mathbb{R}^n)$ by

$$g(x) = \prod_{i=1}^n g_i(x_i').$$

Then

$$\int g \, dx \leq \prod_{i=1}^n \|g_i\|_{n-1}. \tag{3.9}$$

Before proving the theorem, we consider what it says in more detail. If $n = 2$, the theorem states that

$$\int g_1(x_2) g_2(x_1) \, dx_1 dx_2 \leq \left(\int g_1(x_2) \, dx_2\right) \left(\int g_2(x_1) \, dx_1\right),$$

which follows immediately from Fubini's theorem. If $n = 3$, the theorem states that

$$\int g_1(x_2, x_3) g_2(x_1, x_3) g_3(x_1, x_2) \, dx_1 dx_2 dx_3$$
$$\leq \left(\int g_1^2(x_2, x_3) \, dx_2 dx_3\right)^{1/2} \left(\int g_2^2(x_1, x_3) \, dx_1 dx_3\right)^{1/2} \left(\int g_3^2(x_1, x_2) \, dx_1 dx_2\right)^{1/2}.$$

To prove the inequality in this case, we fix x_1 and apply the Cauchy-Schwartz inequality to the x_2x_3-integral of $g_1 \cdot g_2 g_3$. We then use the inequality for $n = 2$ to estimate the x_2x_3-integral of $g_2 g_3$, and integrate the result over x_1. An analogous approach works for higher n.

Note that under the scaling $g_i \mapsto \lambda g_i$, both sides of (3.9) scale in the same way,

$$\int g \, dx \mapsto \left(\prod_{i=1}^n \lambda_i\right) \int g \, dx, \qquad \prod_{i=1}^n \|g_i\|_{n-1} \mapsto \left(\prod_{i=1}^n \lambda_i\right) \prod_{i=1}^n \|g_i\|_{n-1}$$

as must be true for any inequality involving norms. Also, under the spatial rescaling $x \mapsto \lambda x$, we have

$$\int g \, dx \mapsto \lambda^{-n} \int g \, dx,$$

while $\|g_i\|_p \mapsto \lambda^{-(n-1)/p} \|g_i\|_p$, so

$$\prod_{i=1}^n \|g_i\|_p \mapsto \lambda^{-n(n-1)/p} \prod_{i=1}^n \|g_i\|_p$$

Thus, if $p = n - 1$ the two terms scale in the same way, which explains the appearance of the L^{n-1}-norms of the g_i's on the right hand side of (3.9).

PROOF. We use proof by induction. The result is true when $n = 2$. Suppose that it is true for $n-1$ where $n \geq 3$.

For $1 \leq i \leq n$, let $g_i : \mathbb{R}^{n-1} \to \mathbb{R}$ and $g\ : \mathbb{R}^n \to \mathbb{R}$ be the functions given in the theorem. Fix $x_1 \in \mathbb{R}$ and define $g_{x_1} : \mathbb{R}^{n-1} \to \mathbb{R}$ by

$$g_{x_1}(x_1') = g(x_1, x_1').$$

For $2 \leq i \leq n$, let $x_i' = \left(x_1, x_{1,i}'\right)$ where

$$x_{1,i}' = (\hat{x}_1, \dots, \hat{x}_i, \dots x_n) \in \mathbb{R}^{n-2}.$$

Define $g_{i,x_1} : \mathbb{R}^{n-2} \to \mathbb{R}$ and $\tilde{g}_{i,x_1} : \mathbb{R}^{n-1} \to \mathbb{R}$ by

$$g_{i,x_1}\left(x_{1,i}'\right) = g_i\left(x_1, x_{1,i}'\right).$$

Then

$$g_{x_1}(x_1') = g_1(x_1') \prod_{i=2}^{n} g_{i,x_1}\left(x_{1,i}'\right).$$

Using Hölder's inequality with $q = n-1$ and $q' = (n-1)/(n-2)$, we get

$$\begin{aligned}
\int g_{x_1}\, dx_1' &= \int g_1 \left(\prod_{i=2}^{n} g_{i,x_1}\left(x_{1,i}'\right) \right) dx_1' \\
&\leq \|g_1\|_{n-1} \left[\int \left(\prod_{i=2}^{n} g_{i,x_1}\left(x_{1,i}'\right) \right)^{(n-1)/(n-2)} dx_1' \right]^{(n-2)/(n-1)}.
\end{aligned}$$

The induction hypothesis implies that

$$\begin{aligned}
\int \left(\prod_{i=2}^{n} g_{i,x_1}\left(x_{1,i}'\right) \right)^{(n-1)/(n-2)} dx_1' &\leq \prod_{i=2}^{n} \left\| g_{i,x_1}^{(n-1)/(n-2)} \right\|_{n-2} \\
&\leq \prod_{i=2}^{n} \|g_{i,x_1}\|_{n-1}^{(n-1)/(n-2)}.
\end{aligned}$$

Hence,

$$\int g_{x_1}\, dx_1' \leq \|g_1\|_{n-1} \prod_{i=2}^{n} \|g_{i,x_1}\|_{n-1}.$$

Integrating this equation over x_1 and using the generalized Hölder inequality with $p_2 = p_3 = \dots = p_n = n-1$, we get

$$\begin{aligned}
\int g\, dx &\leq \|g_1\|_{n-1} \int \left(\prod_{i=2}^{n} \|g_{i,x_1}\|_{n-1} \right) dx_1 \\
&\leq \|g_1\|_{n-1} \left(\prod_{i=2}^{n} \int \|g_{i,x_1}\|_{n-1}^{n-1}\, dx_1 \right)^{1/(n-1)}.
\end{aligned}$$

Thus, since

$$\begin{aligned}
\int \|g_{i,x_1}\|_{n-1}^{n-1}\, dx_1 &= \int \left(\int \left| g_{i,x_1}(x_{1,i}') \right|^{n-1} dx_{1,i}' \right) dx_1 \\
&= \int \left| g_i(x_i') \right|^{n-1} dx_i' \\
&= \|g_i\|_{n-1}^{n-1},
\end{aligned}$$

we find that

$$\int g\, dx \leq \prod_{i=1}^{n} \|g_i\|_{n-1}.$$

The result follows by induction. □

We now prove the main result.

Theorem 3.28. *Let $1 \leq p < n$, where $n \geq 2$, and let p^* be the Sobolev conjugate of p given in Definition 3.25. Then*

$$\|f\|_{p^*} \leq C\, \|Df\|_p\,, \qquad \textit{for all } f \in C_c^\infty(\mathbb{R}^n)$$

where

$$C(n,p) = \frac{p}{2n}\left(\frac{n-1}{n-p}\right). \tag{3.10}$$

Proof. First, we prove the result for $p = 1$. For $1 \leq i \leq n$, we have

$$|f(x)| \leq \frac{1}{2}\int |\partial_i f(t, x_i')|\, dt.$$

Multiplying these inequalities and taking the $(n-1)$th root, we get

$$|f|^{n/(n-1)} \leq \frac{1}{2^{n/(n-1)}}\, g, \qquad g = \prod_{i=1}^{n} \tilde{g}_i$$

where $\tilde{g}_i(x) = g_i(x_i')$ with

$$g_i(x_i') = \left(\int |\partial_i f(t, x_i')|\, dt\right)^{1/(n-1)}.$$

Theorem 3.27 implies that

$$\int g\, dx \leq \prod_{i=1}^{n} \|g_i\|_{n-1}.$$

Since

$$\|g_i\|_{n-1} = \left(\int |\partial_i f|\, dx\right)^{1/(n-1)}$$

it follows that

$$\int |f|^{n/(n-1)}\, dx \leq \frac{1}{2^{n/(n-1)}}\left(\prod_{i=1}^{n}\int |\partial_i f|\, dx\right)^{1/(n-1)}.$$

Note that $n/(n-1) = 1^*$ is the Sobolev conjugate of 1.

Using the arithmetic-geometric mean inequality,

$$\left(\prod_{i=1}^{n} a_i\right)^{1/n} \leq \frac{1}{n}\sum_{i=1}^{n} a_i,$$

we get

$$\int |f|^{n/(n-1)}\, dx \leq \left(\frac{1}{2n}\sum_{i=1}^{n}\int |\partial_i f|\, dx\right)^{n/(n-1)},$$

or

$$\|f\|_{1^*} \leq \frac{1}{2n}\, \|Df\|_1\,,$$

which proves the result when $p = 1$.

Next suppose that $1 < p < n$. For any $s > 1$, we have

$$\frac{d}{dx}|x|^s = s \operatorname{sgn} x |x|^{s-1}.$$

Thus,

$$\begin{aligned} |f(x)|^s &= \int_{-\infty}^{x_i} \partial_i \left|f(t, x_i')\right|^s \, dt \\ &= s \int_{-\infty}^{x_i} \left|f(t, x_i')\right|^{s-1} \operatorname{sgn}\left[f(t, x_i')\right] \partial_i f(t, x_i') \, dt. \end{aligned}$$

Using Lemma 3.26, it follows that

$$|f(x)|^s \leq \frac{s}{2} \int_{-\infty}^{\infty} \left|f^{s-1}(t, x_i')\partial_i f(t, x_i')\right| \, dt,$$

and multiplication of these inequalities gives

$$|f(x)|^{sn} \leq \left(\frac{s}{2}\right)^n \prod_{i=1}^n \int_{-\infty}^{\infty} \left|f^{s-1}(t, x_i')\partial_i f(t, x_i')\right| \, dt.$$

Applying Theorem 3.27 with the functions

$$g_i(x_i') = \left[\int_{-\infty}^{\infty} \left|f^{s-1}(t, x_i')\partial_i f(t, x_i')\right| \, dt\right]^{1/(n-1)}$$

we find that

$$\|f\|_{sn/(n-1)}^{sn} \leq \frac{s}{2} \prod_{i=1}^n \left\|f^{s-1}\partial_i f\right\|_1.$$

From Hölder's inequality,

$$\left\|f^{s-1}\partial_i f\right\|_1 \leq \left\|f^{s-1}\right\|_{p'} \|\partial_i f\|_p.$$

We have

$$\left\|f^{s-1}\right\|_{p'} = \|f\|_{p'(s-1)}^{s-1}$$

We choose $s > 1$ so that

$$p'(s-1) = \frac{sn}{n-1},$$

which holds if

$$s = p\left(\frac{n-1}{n-p}\right), \qquad \frac{sn}{n-1} = p^*.$$

Then

$$\|f\|_{p^*} \leq \frac{s}{2}\left(\prod_{i=1}^n \|\partial_i f\|_p\right)^{1/n}.$$

Using the arithmetic-geometric mean inequality, we get

$$\|f\|_{p^*} \leq \frac{s}{2n}\left(\sum_{i=1}^n \|\partial_i f\|_p^p\right)^{1/p},$$

which proves the result. □

We can interpret this result roughly as follows: Differentiation of a function increases the strength of its local singularities and improves its decay at infinity. Thus, if $Df \in L^p$, it is reasonable to expect that $f \in L^{p^*}$ for some $p^* > p$ since L^{p^*}-functions have weaker singularities and can decay more slowly at infinity than L^p-functions.

Example 3.29. For $a > 0$, let $f_a : \mathbb{R}^n \to \mathbb{R}$ be the function

$$f_a(x) = \frac{1}{|x|^a}$$

considered in Example 3.7. This function does not belong to $L^q(\mathbb{R}^n)$ for any a since the integral at infinity diverges whenever the integral at zero converges. Let ϕ be a smooth cut-off function that is equal to one for $|x| \le 1$ and zero for $|x| \ge 2$. Then $g_a = \phi f_a$ is an unbounded function with compact support. We have $g_a \in L^q(\mathbb{R}^n)$ if $aq < n$, and $Dg_a \in L^p(\mathbb{R}^n)$ if $p(a+1) < n$ or $ap^* < n$. Thus if $Dg_a \in L^p(\mathbb{R}^n)$, then $g_a \in L^q(\mathbb{R}^n)$ for $1 \le q \le p^*$. On the other hand, the function $h_a = (1-\phi)f_a$ is smooth and decays like $|x|^{-a}$ as $x \to \infty$. We have $h_a \in L^q(\mathbb{R}^n)$ if $qa > n$ and $Dh_a \in L^p(\mathbb{R}^n)$ if $p(a+1) > n$ or $p^*a > n$. Thus, if $Dh_a \in L^p(\mathbb{R}^n)$, then $f \in L^q(\mathbb{R}^n)$ for $p^* \le q < \infty$. The function $f_{ab} = g_a + h_b$ belongs to $L^{p^*}(\mathbb{R}^n)$ for any choice of $a, b > 0$ such that $Df_{ab} \in L^p(\mathbb{R}^n)$. On the other hand, for any $1 \le q \le \infty$ such that $q \ne p^*$, there is a choice of $a, b > 0$ such that $Df_{ab} \in L^p(\mathbb{R}^n)$ but $f_{ab} \notin L^q(\mathbb{R}^n)$.

The constant in Theorem 3.28 is not optimal. For $p = 1$, the best constant is

$$C(n,1) = \frac{1}{n\alpha_n^{1/n}}$$

where α_n is the volume of the unit ball, or

$$C(n,1) = \frac{1}{n\sqrt{\pi}}\left[\Gamma\left(1 + \frac{n}{2}\right)\right]^{1/n}$$

where Γ is the Γ-function. Equality is obtained in the limit of functions that approach the characteristic function of a ball. This result for the best Sobolev constant is equivalent to the isoperimetric inequality that a sphere has minimal area among all surfaces enclosing a given volume.

For $1 < p < n$, the best constant is (Talenti, 1976)

$$C(n,p) = \frac{1}{n^{1/p}\sqrt{\pi}}\left(\frac{p-1}{n-p}\right)^{1-1/p}\left[\frac{\Gamma(1+n/2)\Gamma(n)}{\Gamma(n/p)\Gamma(1+n-n/p)}\right]^{1/n}.$$

Equality holds for functions of the form

$$f(x) = \left(a + b|x|^{p/(p-1)}\right)^{1-n/p}$$

where a, b are positive constants.

The Sobolev inequality in Theorem 3.28 does not hold in the limiting case $p \to n$, $p^* \to \infty$.

Example 3.30. If $\phi(x)$ is a smooth cut-off function that is equal to one for $|x| \le 1$ and zero for $|x| \ge 2$, and

$$f(x) = \phi(x)\log\log\left(1 + \frac{1}{|x|}\right),$$

then $Df \in L^n(\mathbb{R}^n)$, and $f \in W^{1,n}(\mathbb{R})$, but $f \notin L^\infty(\mathbb{R}^n)$.

We can use the Sobolev inequality to prove various embedding theorems. In general, we say that a Banach space X is continuously embedded, or embedded for short, in a Banach space Y if there is a one-to-one, bounded linear map $\imath : X \to Y$. We often think of $\imath$ as identifying elements of the smaller space X with elements of the larger space Y; if X is a subset of Y, then $\imath$ is the inclusion map. The boundedness of $\imath$ means that there is a constant C such that $\|\imath x\|_Y \le C\|x\|_X$ for all $x \in X$, so the weaker Y-norm of $\imath x$ is controlled by the stronger X-norm of x.

We write an embedding as $X \hookrightarrow Y$, or as $X \subset Y$ when the boundedness is understood.

Theorem 3.31. *Suppose that $1 \le p < n$ and $p \le q \le p^*$ where p^* is the Sobolev conjugate of p. Then $W^{1,p}(\mathbb{R}^n) \hookrightarrow L^q(\mathbb{R}^n)$ and*

$$\|f\|_q \le C\|f\|_{W^{1,p}} \qquad \text{for all } f \in W^{1,p}(\mathbb{R}^n)$$

for some constant $C = C(n, p, q)$.

Proof. If $f \in W^{1,p}(\mathbb{R}^n)$, then by Theorem 3.24 there is a sequence of functions $f_n \in C_c^\infty(\mathbb{R}^n)$ that converges to f in $W^{1,p}(\mathbb{R}^n)$. Theorem 3.28 implies that $f_n \to f$ in $L^{p^*}(\mathbb{R}^n)$. In detail: $\{Df_n\}$ converges to Df in L^p so it is Cauchy in L^p; since

$$\|f_n - f_m\|_{p^*} \le C\|Df_n - Df_m\|_p$$

$\{f_n\}$ is Cauchy in L^{p^*}; therefore $f_n \to \tilde{f}$ for some $\tilde{f} \in L^{p^*}$ since L^{p^*} is complete; and $\tilde{f}$ is equivalent to f since a subsequence of $\{f_n\}$ converges pointwise a.e. to $\tilde{f}$, from the L^{p^*} convergence, and to f, from the L^p-convergence.

Thus, $f \in L^{p^*}(\mathbb{R}^n)$ and

$$\|f\|_{p^*} \le C\|Df\|_p.$$

Since $f \in L^p(\mathbb{R}^n)$, Lemma 1.11 implies that for $p < q < p^*$

$$\|f\|_q \le \|f\|_p^\theta \|f\|_{p^*}^{1-\theta}$$

where $0 < \theta < 1$ is defined by

$$\frac{1}{q} = \frac{\theta}{p} + \frac{1-\theta}{p^*}.$$

Therefore, using Theorem 3.28 and the inequality

$$a^\theta b^{1-\theta} \le \left[\theta^\theta (1-\theta)^{1-\theta}\right]^{1/p} (a^p + b^p)^{1/p},$$

we get

$$\begin{aligned}
\|f\|_q &\le C^{1-\theta}\|f\|_p^\theta \|Df\|_p^{1-\theta} \\
&\le C^{1-\theta}\left[\theta^\theta (1-\theta)^{1-\theta}\right]^{1/p} \left(\|f\|_p^p + \|Df\|_p^p\right)^{1/p} \\
&\le C^{1-\theta}\left[\theta^\theta (1-\theta)^{1-\theta}\right]^{1/p} \|f\|_{W^{1,p}}.
\end{aligned}$$

□

Sobolev embedding gives a stronger conclusion for sets Ω with finite measure. In that case, $L^{p^*}(\Omega) \hookrightarrow L^q(\Omega)$ for every $1 \le q \le p^*$, so $W^{1,p}(\Omega) \hookrightarrow L^q(\Omega)$ for $1 \le q \le p^*$, not just $p \le q \le p^*$.

Theorem 3.28 does not, of course, imply that $f \in L^{p^*}(\mathbb{R}^n)$ whenever $Df \in L^p(\mathbb{R}^n)$, since constant functions have zero derivative. To ensure that $f \in L^{p^*}(\mathbb{R}^n)$, we also need to impose a decay condition on f that eliminates the constant functions. In Theorem 3.31, this is provided by the assumption that $f \in L^p(\mathbb{R}^n)$ in addition to $Df \in L^p(\mathbb{R}^n)$. The weakest decay condition we can impose is the following one.

Definition 3.32. A Lebesgue measurable function $f : \mathbb{R}^n \to \mathbb{R}$ vanishes at infinity if for every $\epsilon > 0$ the set $\{x \in \mathbb{R}^n : |f(x)| > \epsilon\}$ has finite Lebesgue measure.

If $f \in L^p(\mathbb{R}^n)$ for some $1 \leq p < \infty$, then f vanishes at infinity. Note that this does not imply that $\lim_{|x|\to\infty} f(x) = 0$.

Example 3.33. Define $f : \mathbb{R} \to \mathbb{R}$ by

$$f = \sum_{n\in\mathbb{N}} \chi_{I_n}, \qquad I_n = \left[n, n + \frac{1}{n^2}\right]$$

where χ_I is the characteristic function of the interval I. Then

$$\int f\, dx = \sum_{n\in\mathbb{N}} \frac{1}{n^2} < \infty,$$

so $f \in L^1(\mathbb{R})$. The limit of $f(x)$ as $|x| \to \infty$ does not exist since $f(x)$ takes on the values 0 and 1 for arbitrarily large values of x. Nevertheless, f vanishes at infinity since for any $\epsilon < 1$,

$$|\{x \in \mathbb{R} : |f(x)| > \epsilon\}| = \sum_{n\in\mathbb{N}} \frac{1}{n^2},$$

which is finite.

Example 3.34. The function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 1/\log x & \text{if } x \geq 2 \\ 0 & \text{if } x < 2 \end{cases}$$

vanishes at infinity, but $f \notin L^p(\mathbb{R})$ for any $1 \leq p < \infty$.

The Sobolev embedding theorem remains true for functions that vanish at infinity.

Theorem 3.35. *Suppose that $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ is weakly differentiable with $Df \in L^p(\mathbb{R}^n)$ where $1 \leq p < n$ and f vanishes at infinity. Then $f \in L^{p^*}(\mathbb{R}^n)$ and*

$$\|f\|_{p^*} \leq C\|Df\|_p$$

where C is given in (3.10).

As before, we prove this by approximating f with smooth compactly supported functions. We omit the details.

3.8. Sobolev embedding: $p > n$

> Friedrichs was a great lover of inequalities, and that affected me very much. The point of view was that the inequalities are more interesting than the equalities, the identities.[3]

[3] Louis Nirenberg on K. O. Friedrichs, from *Notices of the AMS*, April 2002.

In the previous section, we saw that if the weak derivative of a function that vanishes at infinity belongs to $L^p(\mathbb{R}^n)$ with $p < n$, then the function has improved integrability properties and belongs to $L^{p^*}(\mathbb{R}^n)$. Even though the function is weakly differentiable, it need not be continuous. In this section, we show that if the derivative belongs to $L^p(\mathbb{R}^n)$ with $p > n$ then the function (or a pointwise a.e. equivalent version of it) is continuous, and in fact Hölder continuous. The following result is due to Morrey (1940). The main idea is to estimate the difference $|f(x) - f(y)|$ in terms of Df by the mean value theorem, average the result over a ball $B_r(x)$ and estimate the result in terms of $\|Df\|_p$ by Hölder's inequality.

Theorem 3.36. *Let $n < p < \infty$ and*

$$\alpha = 1 - \frac{n}{p},$$

with $\alpha = 1$ if $p = \infty$. Then there are constants $C = C(n,p)$ such that

$$[f]_\alpha \leq C \|Df\|_p \qquad \text{for all } f \in C_c^\infty(\mathbb{R}^n), \tag{3.11}$$

$$\sup_{\mathbb{R}^n} |f| \leq C \|f\|_{W^{1,p}} \qquad \text{for all } f \in C_c^\infty(\mathbb{R}^n), \tag{3.12}$$

where $[\cdot]_\alpha$ denotes the Hölder seminorm $[\cdot]_{\alpha,\mathbb{R}^n}$ defined in (1.1).

Proof. First we prove that there exists a constant C depending only on n such that for any ball $B_r(x)$

$$⨍_{B_r(x)} |f(x) - f(y)|\, dy \leq C \int_{B_r(x)} \frac{|Df(y)|}{|x-y|^{n-1}}\, dy \tag{3.13}$$

Let $w \in \partial B_1(0)$ be a unit vector. For $s > 0$

$$f(x+sw) - f(x) = \int_0^s \frac{d}{dt} f(x+tw)\, dt = \int_0^s Df(x+tw)\cdot w\, dt,$$

and therefore since $|w| = 1$

$$|f(x+sw) - f(x)| \leq \int_0^s |Df(x+tw)|\, dt.$$

Integrating this inequality with respect to w over the unit sphere, we get

$$\int_{\partial B_1(0)} |f(x) - f(x+sw)|\, dS(w) \leq \int_{\partial B_1(0)} \left(\int_0^s |Df(x+tw)|\, dt\right) dS(w).$$

From Proposition 1.45,

$$\begin{aligned}\int_{\partial B_1(0)} \left(\int_0^s |Df(x+tw)|\, dt\right) dS(w) &= \int_{\partial B_1(0)} \int_0^s \frac{|Df(x+tw)|}{t^{n-1}}\, t^{n-1}\, dt dS(w)\\ &= \int_{B_s(x)} \frac{|Df(y)|}{|x-y|^{n-1}}\, dy,\end{aligned}$$

Thus,

$$\int_{\partial B_1(0)} |f(x) - f(x+sw)|\, dS(w) \leq \int_{B_s(x)} \frac{|Df(y)|}{|x-y|^{n-1}}\, dy.$$

Using Proposition 1.45 together with this inequality, and estimating the integral over $B_s(x)$ by the integral over $B_r(x)$ for $s \leq r$, we find that

$$\begin{aligned}
\int_{B_r(x)} |f(x) - f(y)|\, dy &= \int_0^r \left(\int_{\partial B_1(0)} |f(x) - f(x + sw)|\, dS(w) \right) s^{n-1}\, ds \\
&\leq \int_0^r \left(\int_{B_s(x)} \frac{|Df(y)|}{|x-y|^{n-1}}\, dy \right) s^{n-1}\, ds \\
&\leq \left(\int_0^r s^{n-1}\, ds \right) \left(\int_{B_r(x)} \frac{|Df(y)|}{|x-y|^{n-1}}\, dy \right) \\
&\leq \frac{r^n}{n} \int_{B_r(x)} \frac{|Df(y)|}{|x-y|^{n-1}}\, dy
\end{aligned}$$

This gives (3.13) with $C = (n\alpha_n)^{-1}$.

Next, we prove (3.11). Suppose that $x, y \in \mathbb{R}^n$. Let $r = |x - y|$ and $\Omega = B_r(x) \cap B_r(y)$. Then averaging the inequality

$$|f(x) - f(y)| \leq |f(x) - f(z)| + |f(y) - f(z)|$$

with respect to z over Ω, we get

$$|f(x) - f(y)| \leq \fint_\Omega |f(x) - f(z)|\, dz + \fint_\Omega |f(y) - f(z)|\, dz. \tag{3.14}$$

From (3.13) and Hölder's inequality,

$$\begin{aligned}
\fint_\Omega |f(x) - f(z)|\, dz &\leq \fint_{B_r(x)} |f(x) - f(z)|\, dz \\
&\leq C \int_{B_r(x)} \frac{|Df(y)|}{|x-y|^{n-1}}\, dy \\
&\leq C \left(\int_{B_r(x)} |Df|^p\, dz \right)^{1/p} \left(\int_{B_r(x)} \frac{dz}{|x-z|^{p'(n-1)}} \right)^{1/p'}.
\end{aligned}$$

We have

$$\left(\int_{B_r(x)} \frac{dz}{|x-z|^{p'(n-1)}} \right)^{1/p'} = C \left(\int_0^r \frac{r^{n-1} dr}{r^{p'(n-1)}} \right)^{1/p'} = C r^{1-n/p}$$

where C denotes a generic constant depending on n and p. Thus,

$$\fint_\Omega |f(x) - f(z)|\, dz \leq C r^{1-n/p} \|Df\|_{L^p(\mathbb{R}^n)},$$

with a similar estimate for the integral in which x is replaced by y. Using these estimates in (3.14) and setting $r = |x - y|$, we get

$$|f(x) - f(y)| \leq C|x - y|^{1-n/p} \|Df\|_{L^p(\mathbb{R}^n)}, \tag{3.15}$$

which proves (3.11).

Finally, we prove (3.12). For any $x \in \mathbb{R}^n$, using (3.15), we find that

$$\begin{aligned} |f(x)| &\leq \fint_{B_1(x)} |f(x) - f(y)|\, dy + \fint_{B_1(x)} |f(y)|\, dy \\ &\leq C \|Df\|_{L^p(\mathbb{R}^n)} + C \|f\|_{L^p(B_1(x))} \\ &\leq C \|f\|_{W^{1,p}(\mathbb{R}^n)}, \end{aligned}$$

and taking the supremum with respect to x, we get (3.12). □

Combining these estimates for

$$\|f\|_{C^{0,\alpha}} = \sup |f| + [f]_\alpha$$

and using a density argument, we get the following theorem. We denote by $C_0^{0,\alpha}(\mathbb{R}^n)$ the space of Hölder continuous functions f whose limit as $x \to \infty$ is zero, meaning that for every $\epsilon > 0$ there exists a compact set $K \subset \mathbb{R}^n$ such that $|f(x)| < \epsilon$ if $x \in \mathbb{R}^n \setminus K$.

Theorem 3.37. *Let $n < p < \infty$ and $\alpha = 1 - n/p$. Then*

$$W^{1,p}(\mathbb{R}^n) \hookrightarrow C_0^{0,\alpha}(\mathbb{R}^n)$$

and there is a constant $C = C(n,p)$ such that

$$\|f\|_{C^{0,\alpha}} \leq C \|f\|_{W^{1,p}} \qquad \text{for all } f \in C_c^\infty(\mathbb{R}^n).$$

Proof. From Theorem 3.24, the mollified functions $\eta^\epsilon * f^\epsilon \to f$ in $W^{1,p}(\mathbb{R}^n)$ as $\epsilon \to 0^+$, and by Theorem 3.36

$$|f^\epsilon(x) - f^\epsilon(y)| \leq C|x-y|^{1-n/p} \|Df^\epsilon\|_{L^p}.$$

Letting $\epsilon \to 0^+$, we find that

$$|f(x) - f(y)| \leq C|x-y|^{1-n/p} \|Df\|_{L^p}$$

for all Lebesgue points $x, y \in \mathbb{R}^n$ of f. Since these form a set of measure zero, f extends by uniform continuity to a uniformly continuous function on $\mathbb{R}^n$.

Also from Theorem 3.24, the function $f \in W^{1,p}(\mathbb{R}^n)$ is a limit of compactly supported functions, and from (3.12), f is the uniform limit of compactly supported functions, which implies that its limit as $x \to \infty$ is zero. □

We state two results without proof (see §5.8 of [**8**]).

For $p = \infty$, the same proof as the proof of (3.11), using Hölder's inequality with $p = \infty$ and $p' = 1$, shows that $f \in W^{1,\infty}(\mathbb{R}^n)$ is Lipschitz continuous, with

$$[f]_1 \leq C \|Df\|_{L^\infty}.$$

A function in $W^{1,\infty}(\mathbb{R}^n)$ need not approach zero at infinity. We have in this case the following characterization of Lipschitz functions.

Theorem 3.38. *A function $f \in L^1_{\mathrm{loc}}(\mathbb{R}^n)$ is Lipschitz continuous if and only if it is weakly differentiable and $Df \in L^\infty(\mathbb{R}^n)$.*

When $n < p \leq \infty$, the above estimates can be used to prove that pointwise derivative of a Sobolev function exists almost everywhere and agrees with the weak derivative.

Theorem 3.39. *If $f \in W^{1,p}_{\mathrm{loc}}(\mathbb{R}^n)$ for some $n < p \leq \infty$, then f is differentiable pointwise a.e. and the pointwise derivative coincides with the weak derivative.*

3.9. Boundary values of Sobolev functions

If $f \in C(\overline{\Omega})$ is a continuous function on the closure of a smooth domain Ω, we can define the boundary values of f pointwise as a continuous function on the boundary $\partial\Omega$. We can also do this when Sobolev embedding implies that a function is Hölder continuous. In general, however, a Sobolev function is not equivalent pointwise a.e. to a continuous function and the boundary of a smooth open set has measure zero, so the boundary values cannot be defined pointwise. For example, we cannot make sense of the boundary values of an L^p-function as an L^p-function on the boundary.

Example 3.40. Suppose $T : C^\infty([0,1]) \to \mathbb{R}$ is the map defined by $T : \phi \mapsto \phi(0)$. If $\phi^\epsilon(x) = e^{-x^2/\epsilon}$, then $\|\phi^\epsilon\|_{L^1} \to 0$ as $\epsilon \to 0^+$, but $\phi^\epsilon(0) = 1$ for every $\epsilon > 0$. Thus, T is not bounded (or even closed) and we cannot extend it by continuity to $L^1(0,1)$.

Nevertheless, we can define the boundary values of a Sobolev function at the expense of a loss of smoothness in restricting the function to the boundary. To do this, we show that the linear map on smooth functions that gives their boundary values is bounded with respect to appropriate Sobolev norms. We then extend the map by continuity to Sobolev functions, and the resulting trace map defines their boundary values.

We consider the basic case of a half-space $\mathbb{R}^n_+$. We write $x = (x', x_n) \in \mathbb{R}^n_+$ where $x_n > 0$ and $(x', 0) \in \partial\mathbb{R}^n_+ = \mathbb{R}^{n-1}$.

The Sobolev space $W^{1,p}(\mathbb{R}^n_+)$ consists of functions $f \in L^p(\mathbb{R}^n_+)$ that are weakly differentiable in $\mathbb{R}^n_+$ with $Df \in L^p(\mathbb{R}^n_+)$. We begin with a result which states that we can extend functions $f \in W^{1,p}(\mathbb{R}^n_+)$ to functions in $W^{1,p}(\mathbb{R}^n)$ without increasing their norm. An extension may be constructed by reflecting a function across the boundary $\partial\mathbb{R}^n_+$ in a way that preserves its differentiability. Such an extension map E is not, of course, unique.

Theorem 3.41. *There is a bounded linear map*

$$E : W^{1,p}(\mathbb{R}^n_+) \to W^{1,p}(\mathbb{R}^n)$$

such that $Ef = f$ pointwise a.e. in $\mathbb{R}^n_+$ and for some constant $C = C(n,p)$

$$\|Ef\|_{W^{1,p}(\mathbb{R}^n)} \leq C \|f\|_{W^{1,p}(\mathbb{R}^n_+)} .$$

The following approximation result may be proved by extending a Sobolev function from $\mathbb{R}^n_+$ to $\mathbb{R}^n$, mollifying the extension, and restricting the result to the half-space.

Theorem 3.42. *The space $C_c^\infty(\overline{\mathbb{R}}^n_+)$ of smooth functions is dense in $W^{k,p}(\mathbb{R}^n_+)$.*

Functions $f : \overline{\mathbb{R}}^n_+ \to \mathbb{R}$ in $C_c^\infty(\overline{\mathbb{R}}^n_+)$ need not vanish on the boundary $\partial\mathbb{R}^n_+$. On the other hand, functions in the space $C_c^\infty(\mathbb{R}^n_+)$ of smooth functions whose support is contained in the open half space $\mathbb{R}^n_+$ do vanish on the boundary, and it is not true that this space is dense in $W^{k,p}(\mathbb{R}^n_+)$. Roughly speaking, we can only approximate by functions in $C_c^\infty(\mathbb{R}^n_+)$ Sobolev functions that 'vanish on the boundary'. We make the following definition.

Definition 3.43. The space $W_0^{k,p}(\mathbb{R}^n_+)$ is the closure of $C_c^\infty(\mathbb{R}^n_+)$ in $W^{k,p}(\mathbb{R}^n_+)$.

The interpretation of $W_0^{1,p}(\mathbb{R}_+^n)$ as the space of Sobolev functions that vanish on the boundary is made more precise in the following theorem, which shows the existence of a trace map T that maps a Sobolev function to its boundary values, and states that functions in $W_0^{1,p}(\mathbb{R}_+^n)$ are the ones whose trace is equal to zero.

Theorem 3.44. *For $1 \le p < \infty$, there is a bounded linear operator*

$$T : W^{1,p}(\mathbb{R}_+^n) \to L^p(\partial\mathbb{R}_+^n)$$

such that for any $f \in C_c^\infty(\overline{\mathbb{R}}_+^n)$

$$(Tf)(x') = f(x', 0)$$

and

$$\|Tf\|_{L^p(\mathbb{R}^{n-1})} \le C \|f\|_{W^{1,p}(\mathbb{R}_+^n)}$$

for some constant C depending only on p. Furthermore, $f \in W_0^{k,p}(\mathbb{R}_+^n)$ if and only if $Tf = 0$.

Proof. First, we consider $f \in C_c^\infty(\overline{\mathbb{R}}_+^n)$. For $x' \in \mathbb{R}^{n-1}$ and $p \ge 1$, we have

$$|f(x', 0)|^p \le p \int_0^\infty |f(x', t)|^{p-1} |\partial_n f(x', t)| \, dt.$$

Hence, using Hölder's inequality and the identity $p'(p-1) = p$, we get

$$\begin{aligned}
\int |f(x', 0)|^p \, dx' &\le p \int_0^\infty |f(x', t)|^{p-1} |\partial_n f(x', t)| \, dx'dt \\
&\le p \left(\int_0^\infty |f(x', t)|^{p'(p-1)} \, dx'dt \right)^{1/p'} \left(\int_0^\infty |\partial_n f(x', t)|^p \, dx'dt \right)^{1/p} \\
&\le p \|f\|_p^{p-1} \|\partial_n f\|_p \\
&\le p\|f\|_{W^{k,p}}^p.
\end{aligned}$$

The trace map

$$T : C_c^\infty(\overline{\mathbb{R}}_+^n) \to C_c^\infty(\mathbb{R}^{n-1})$$

is therefore bounded with respect to the $W^{1,p}(\mathbb{R}_+^n)$ and $L^p(\partial\mathbb{R}_+^n)$ norms, and extends by density and continuity to a map between these spaces.

It follows immediately that $Tf = 0$ if $f \in W_0^{k,p}(\mathbb{R}_+^n)$. We omit the proof that $Tf = 0$ implies that $f \in W_0^{k,p}(\mathbb{R}_+^n)$ (see [**8**]). □

If $p = 1$, the trace $T : W^{1,1}(\mathbb{R}_+^n) \to L^1(\mathbb{R}^{n-1})$ is onto, but if $1 < p < \infty$ the range of T is not all of L^p. In that case, $T : W^{1,p}(\mathbb{R}_+^n) \to B^{1-1/p,p}(\mathbb{R}^{n-1})$ maps $W^{1,p}$ onto a Besov space $B^{1-1/p,p}$; roughly speaking, this is a Sobolev space of functions with fractional derivatives, and there is a loss of $1/p$ derivatives in restricting a function to the boundary [**20**].

Note that if $f \in W_0^{2,p}(\mathbb{R}_+^n)$, then $\partial_i f \in W_0^{1,p}(\mathbb{R}_+^n)$, so $T(\partial_i f) = 0$. Thus, both f and Df vanish on the boundary. The correct way to formulate the condition that f has weak derivatives of order less than or equal to two and satisfies the Dirichlet condition $f = 0$ on the boundary is that $f \in W^{2,p}(\mathbb{R}_+^n) \cap W_0^{1,p}(\mathbb{R}_+^n)$.

3.10. Compactness results

A Banach space X is compactly embedded in a Banach space Y, written $X \Subset Y$, if the embedding $\imath : X \to Y$ is compact. That is, $\imath$ maps bounded sets in X to precompact sets in Y; or, equivalently, if $\{x_n\}$ is a bounded sequence in X, then $\{\imath x_n\}$ has a convergent subsequence in Y.

An important property of the Sobolev embeddings is that they are compact on domains with finite measure. This corresponds to the rough principle that uniform bounds on higher derivatives imply compactness with respect to lower derivatives. The compactness of the Sobolev embeddings, due to Rellich and Kondrachov, depend on the Arzelà-Ascoli theorem. We will prove a version for $W_0^{1,p}(\Omega)$ by use of the L^p-compactness criterion in Theorem 1.15.

Theorem 3.45. *Let Ω be a bounded open set in $\mathbb{R}^n$, $1 \le p < n$, and $1 \le q < p^*$. If $\mathcal{F}$ is a bounded set in $W_0^{1,p}(\Omega)$, then $\mathcal{F}$ is precompact in $L^q(\mathbb{R}^n)$.*

PROOF. By a density argument, we may assume that the functions in $\mathcal{F}$ are smooth and $\operatorname{spt} f \Subset \Omega$. We may then extend the functions and their derivatives by zero to obtain smooth functions on $\mathbb{R}^n$, and prove that $\mathcal{F}$ is precompact in $L^q(\mathbb{R}^n)$.

Condition (1) in Theorem 1.15 follows immediately from the boundedness of Ω and the Sobolev embeddeding theorem: for all $f \in \mathcal{F}$,

$$\|f\|_{L^q(\mathbb{R}^n)} = \|f\|_{L^q(\Omega)} \le C\|f\|_{L^{p^*}(\Omega)} \le C\|Df\|_{L^p(\mathbb{R}^n)} \le C$$

where C denotes a generic constant that does not depend on f. Condition (2) is satisfied automatically since the supports of all functions in $\mathcal{F}$ are contained in the same bounded set.

To verify (3), we first note that since Df is supported inside the bounded open set Ω,

$$\|Df\|_{L^1(\mathbb{R}^n)} \le C\,\|Df\|_{L^p(\mathbb{R}^n)}\,.$$

Fix $h \in \mathbb{R}^n$ and let $f_h(x) = f(x+h)$ denote the translation of f by h. Then

$$|f_h(x) - f(x)| = \left|\int_0^1 h \cdot Df(x+th)\,dt\right| \le |h| \int_0^1 |Df(x+th)|\,dt.$$

Integrating this inequality with respect to x and using Fubini's theorem to exchange the order of integration on the right-hand side, together with the fact that the inner x-integral is independent of t, we get

$$\int_{\mathbb{R}^n} |f_h(x) - f(x)|\,dx \le |h|\,\|Df\|_{L^1(\mathbb{R}^n)} \le C|h|\,\|Df\|_{L^p(\mathbb{R}^n)}\,.$$

Thus,

$$\|f_h - f\|_{L^1(\mathbb{R}^n)} \le C|h|\,\|Df\|_{L^p(\mathbb{R}^n)}\,. \tag{3.16}$$

Using the interpolation inequality in Lemma 1.11, we get for any $1 \le q < p^*$ that

$$\|f_h - f\|_{L^q(\mathbb{R}^n)} \le \|f_h - f\|_{L^1(\mathbb{R}^n)}^{\theta}\,\|f_h - f\|_{L^{p^*}(\mathbb{R}^n)}^{1-\theta} \tag{3.17}$$

where $0 < \theta \le 1$ is given by

$$\frac{1}{q} = \theta + \frac{1-\theta}{p^*}.$$

The Sobolev embedding theorem implies that

$$\|f_h - f\|_{L^{p^*}(\mathbb{R}^n)} \le C\,\|Df\|_{L^p(\mathbb{R}^n)}\,.$$

Using this inequality and (3.16) in (3.17), we get

$$\|f_h - f\|_{L^q(\mathbb{R}^n)} \leq C|h|^\theta \|Df\|_{L^p(\mathbb{R}^n)}.$$

It follows that $\mathcal{F}$ is L^q-equicontinuous if the derivatives of functions in $\mathcal{F}$ are uniformly bounded in L^p, and the result follows. □

Equivalently, this theorem states that if $\{f_. k \in \mathbb{N}\}$ is a sequence of functions in $W_0^{1,p}(\Omega)$ such that

$$\|f_k\|_{W^{1,p}} \leq C \qquad \text{for all } k \in \mathbb{N},$$

for some constant C, then there exists a subsequence f_{k_i} and a function $f \in L^q(\Omega)$ such that

$$f_{k_i} \to f \qquad \text{as } i \to \infty \text{ in } L^q(\Omega).$$

The assumptions that the domain Ω satisfies a boundedness condition and that $q < p^*$ are necessary.

Example 3.46. If $\phi \in W^{1,p}(\mathbb{R}^n)$ and $f_m(x) = \phi(x - c_m)$, where $c_m \to \infty$ as $m \to \infty$, then $\|f_m\|_{W^{1,p}} = \|\phi\|_{W^{1,p}}$ is constant, but $\{f_m\}$ has no convergent subsequence in L^q since the functions 'escape' to infinity. Thus, compactness does not hold without some limitation on the decay of the functions.

Example 3.47. For $1 \leq p < n$, define $f_k : \mathbb{R}^n \to \mathbb{R}$ by

$$f_k(x) = \begin{cases} k^{n/p^*}(1 - k|x|) & \text{if } |x| < 1/k, \\ 0 & \text{if } |x| \geq 1/k. \end{cases}$$

Then $\operatorname{spt} f_k \subset \overline{B}_1(0)$ for every $k \in \mathbb{N}$ and $\{f_k\}$ is bounded in $W^{1,p}(\mathbb{R}^n)$, but no subsequence converges strongly in $L^{p^*}(\mathbb{R}^n)$.

The loss of compactness in the critical case $q = p^*$ has received a great deal of study (for example, in the concentration compactness principle of P.L. Lions).

If Ω is a smooth and bounded domain, the use of an extension map implies that $W^{1,p}(\Omega) \Subset L^q(\Omega)$. For an example of the loss of this compactness in a bounded domain with an irregular boundary, see [**20**].

Theorem 3.48. *Let Ω be a bounded open set in $\mathbb{R}^n$, and $n < p < \infty$. Suppose that $\mathcal{F}$ is a set of functions whose weak derivative belongs to $L^p(\mathbb{R}^n)$ such that: (a)* $\operatorname{spt} f \Subset \Omega$*; (b) there exists a constant C such that*

$$\|Df\|_{L^p} \leq C \qquad \textit{for all } f \in \mathcal{F}.$$

Then $\mathcal{F}$ is precompact in $C_0(\mathbb{R}^n)$.

Proof. Theorem 3.36 implies that the set $\mathcal{F}$ is bounded and equicontinuous, so the result follows immediately from the Arzelà-Ascoli theorem. □

In other words, if $\{f_m : m \in \mathbb{N}\}$ is a sequence of functions in $W^{1,p}(\mathbb{R}^n)$ such that $\operatorname{spt} f_m \subset \Omega$, where $\Omega \Subset \mathbb{R}^n$, and

$$\|f_m\|_{W^{1,p}} \leq C \qquad \text{for all } m \in \mathbb{N}$$

for some constant C, then there exists a subsequence f_{m_k} such that $f_{n_k} \to f$ uniformly, in which case $f \in C_c(\mathbb{R}^n)$.

3.11. Sobolev functions on $\Omega \subset \mathbb{R}^n$

Here, we briefly outline how ones transfers the results above to Sobolev spaces on domains other than $\mathbb{R}^n$ or $\mathbb{R}^n_+$.

Suppose that Ω is a smooth, bounded domain in $\mathbb{R}^n$. We may cover the closure $\overline{\Omega}$ by a collection of open balls contained in Ω and open balls with center $x \in \partial\Omega$. Since $\overline{\Omega}$ is compact, there is a finite collection $\{B_i : 1 \le i \le N\}$ of such open balls that covers $\overline{\Omega}$. There is a partition of unity $\{\psi_i : 1 \le i \le N\}$ subordinate to this cover consisting of functions $\psi_i \in C_c^\infty(B_i)$ such that $0 \le \psi_i \le 1$ and $\sum_i \psi_i = 1$ on $\overline{\Omega}$.

Given any function $f \in L^1_{\mathrm{loc}}(\Omega)$, we may write $f = \sum_i f_i$ where $f_i = \psi_i f$ has compact support in B_i for balls whose center belongs to Ω, and in $B_i \cap \overline{\Omega}$ for balls whose center belongs to $\partial\Omega$. In these latter balls, we may 'straighten out the boundary' by a smooth map. After this change of variables, we get a function f_i that is compactly supported in $\overline{\mathbb{R}}^n_+$. We may then apply the previous results to the functions $\{f_i : 1 \le i \le N\}$.

Typically, results about $W_0^{k,p}(\Omega)$ do not require assumptions on the smoothness of $\partial\Omega$; but results about $W^{k,p}(\Omega)$ — for example, the existence of a bounded extension operator $E : W^{k,p}(\Omega) \to W^{k,p}(\mathbb{R}^n)$ — only hold if $\partial\Omega$ satisfies an appropriate smoothness or regularity condition *e.g.* a C^k, Lipschitz, segment, or cone condition [**1**].

The statement of the embedding theorem for higher order derivatives extends in a straightforward way from the one for first order derivatives. For example,

$$W^{k,p}(\mathbb{R}^n) \hookrightarrow L^q(\mathbb{R}^n) \qquad \text{if} \quad \frac{1}{q} = \frac{1}{p} - \frac{k}{n}.$$

The result for smooth bounded domains is summarized in the following theorem. As before, $X \subset Y$ denotes a continuous embedding of X into Y, and $X \Subset Y$ denotes a compact embedding.

Theorem 3.49. *Suppose that Ω is a bounded open set in $\mathbb{R}^n$ with C^1 boundary, $k, m \in \mathbb{N}$ with $k \ge m$, and $1 \le p < \infty$.*

(1) *If $kp < n$, then*

$$W^{k,p}(\Omega) \Subset L^q(\Omega) \qquad \text{for } 1 \le q < np/(n-kp);$$
$$W^{k,p}(\Omega) \subset L^q(\Omega) \qquad \text{for } q = np/(n-kp).$$

More generally, if $(k-m)p < n$, then

$$W^{k,p}(\Omega) \Subset W^{m,q}(\Omega) \qquad \text{for } 1 \le q < np/\left(n-(k-m)p\right);$$
$$W^{k,p}(\Omega) \subset W^{m,q}(\Omega) \qquad \text{for } q = np/\left(n-(k-m)p\right).$$

(2) *If $kp = n$, then*

$$W^{k,p}(\Omega) \Subset L^q(\Omega) \qquad \text{for } 1 \le q < \infty.$$

(3) *If $kp > n$, then*

$$W^{k,p}(\Omega) \Subset C^{0,\mu}\left(\overline{\Omega}\right)$$

for $0 < \mu < k - n/p$ if $k - n/p < 1$, for $0 < \mu < 1$ if $k - n/p = 1$, and for $\mu = 1$ if $k - n/p > 1$; and

$$W^{k,p}(\Omega) \subset C^{0,\mu}\left(\overline{\Omega}\right)$$

for $\mu = k - n/p$ *if* $k - n/p < 1$. *More generally, if* $(k-m)p > n$, *then*

$$W^{k,p}(\Omega) \Subset C^{m,\mu}\left(\overline{\Omega}\right)$$

for $0 < \mu < k-m-n/p$ *if* $k-m-n/p < 1$, *for* $0 < \mu < 1$ *if* $k-m-n/p = 1$, *and for* $\mu = 1$ *if* $k - m - n/p > 1$; *and*

$$W^{k,p}(\Omega) \subset C^{m,\mu}\left(\overline{\Omega}\right)$$

for $\mu = k - m - n/p$ *if* $k - m - n/p = 0$.

These results hold for arbitrary bounded open sets Ω if $W^{k,p}(\Omega)$ is replaced by $W_0^{k,p}(\Omega)$.

Example 3.50. If $u \in W^{n,1}(\mathbb{R}^n)$, then $u \in C_0(\mathbb{R}^n)$. This can be seen from the equality

$$u(x) = \int_0^{x_1} \dots \int_0^{x_n} \partial_1 \cdots \partial_n u(x')\, dx_1' \dots dx_n',$$

which holds for all $u \in C_c^\infty(\mathbb{R}^n)$ and a density argument. In general, however, it is not true that $u \in L^\infty$ in the critical case $kp = n$ *c.f.* Example 3.30.